THE IMPACT OF ERROR-PRONE FEEDBACK ON LEARNING IN A MULTIPLE-CUE DECISION PROBLEM

A Thesis

Presented to

The Faculty of the Division of Graduate

Studies and Research

By

Val Michael Miklausich

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

In the School of Industrial Engineering

Georgia Institute of Technology

December, 1973

THE IMPACT OF ERROR-PRONE FEEDBACK ON

LEARNING IN A MULTIPLE-CUE DECISION PROBLEM

APPROVED:

Terry Connolly, Chairman

Norman R. Baker

Douglas C. Montgomery

Date Approved by Chairman 8 Oct 75

ACKNOWLEDGMENTS

Dr. T. Connolly was a tremendous help to me in completing this thesis. A mere "thank you" is not enough for all the time and effort he gave me. To Drs. D. C. Montgomery and N. R. Baker I owe many thanks for their time and many helpful suggestions. I would also like to thank Dr. J. Walker and Mr. Frank Alt for the time and effort they gave me. Also I would like to say "thank you" to my parents, Mr. and Mrs. Miklausich, for the inspiration they gave me throughout my life and to Ginny who makes all my efforts worthwhile.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

INTRODUCTION

The following study was performed in an effort to learn more about decisions made by people. From studies such as Bowman, (1963), and Goldberg, (1968), it is known that certain models predict the real-world outcomes better than other models. The class of decision problems considered here are those that can be analyzed by the lens model.

The lens model was developed by Brunswik, (1952), and a detailed description of it applied to clinical judgements was published by Hammond, (1955). Since the appearance of these two papers the literature dealing with the lens model has expanded rapidly. Because the lens model is a fairly recent development there are many types of experimental areas to be explored. In this study the area of the impact of error in feedback in a multiple-cue decision problem will be explored.

The results of this study indicate that subjects do not predict outcomes in high error in feedback conditions as well as they do for low error in feedback. Also it was found that the degradation of the subjects' ability to predict outcomes was primarily due to the failure of their estimates to fit the linear regression equation of their estimates, rather than to the failure of the linear regression equation estimates of the subjects' predictions of the outcome variable to correlate with the linear regression equation estimates of the outcome variable values. The study makes clear the development and significance of such terms as error in feedback, linear regression, and correlations, as used herein.

CHAPTER I

LENS MODEL

Basic Terminology of the Lens Model

Consider the following simple decision making situation. An individual works all day in a windowless office. Each day just before the end of work, a friend visits him and reports some data about the outside weather conditions. The friend reports such things as temperature, the brightness of the sky and the humidity. The office worker uses these data to decide whether or not he should take his umbrella with him when he leaves. Finally, he receives some feedback on the correctness of his decision when he leaves the office and discovers the actual weather conditions for himself.

Several features of this situation are of interest. First, the decision is based on an estiamte of some aspect of external reality which is not itself directly observed. Second, the data on which the estimate is based are subject to error, in that they are not uneqivocally related to the external state of interest. Finally, the feedback the decision maker receives is itself subject to error, in that the situation he observes may have changed from that generating his data. It will be apparent that these characteristics describe a rather large range of decision making situations.

An approach to the analysis of such situations is provided by Brunswik's lens model. (Brunswik, (1952); Hammond, (1955)). The basic elements of this model are shown in Figure 1.

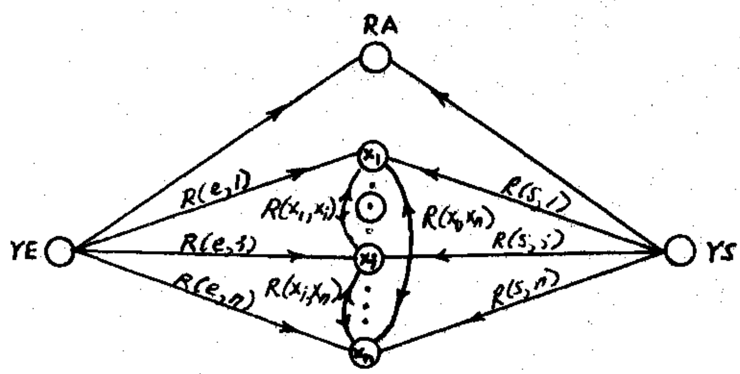


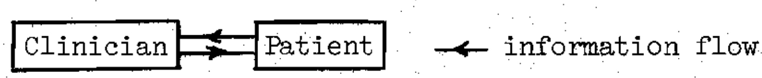
Figure 1. Basic Elements of Lens Model

In the usual terminology for this model, the decision maker is seen as using a set of information-bearing "cues" x_i to arrive at an estimate YS of the value of a "Distal" or "Criterion" variable YE. Each cue is related to the distal variable in some way, the extent of this relationship being indicated by the correlation coefficient R(e, i), termed the "Ecological Validity" of the i^{th} cue. Similarly, the extent of the relationship between each cue and the decision maker's estimate is indicated by the correlation coefficient R(S,i), termed the subject's "Utilization Coefficient" of the i^{th} cue. Each pair of cues will, in general, show some interrelationship, indicated by the cue intercorrelations R(i,j). A simple measure of the decision maker's performance is provided by the correlation R(YE, YS), the "Achievement Index."

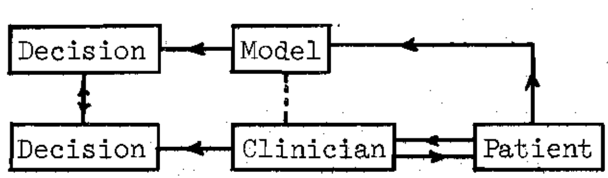
Conceptual Development of the Lens Model

The lens model is concerned with situations in which a person finds himself confronted with different sources of information relevant to some distal variable, which can be used to estimate an outcome. The sources of information will be called cues. For example, cues can be symptoms that occur together and characterize a particular disease. A clinician looking at the symptoms must make some decision. "Exactly how he arrives at a decision is not known even to him." (Hammond, (1955)).

Since a clinician does not know how he arrives at a decision the information flow between clinician and patient will be considered. For example, the following situation can exist.



The above can represent the clinician studying his patient by way of a test or just by examination. Since a clinician is subject to error he should himself be analyzed. For example, the situation can be redrawn as follows:



——— information flow

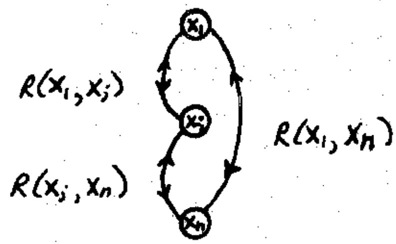
- - - indicates statistics can be used to compare

The difference in this latter reformulation of the situation is that a model of the clinician has been added. The dotted line indicates that the model and clinician can be compared by the use of statistics. The model used will be the lens model but it must be remembered that we are studying decision problems where the lens model is the appropriate model to use. For some other decision problems, (Bowman, (1963)), the lens model would not be the appropriate model.

Formulation of the Lens Model

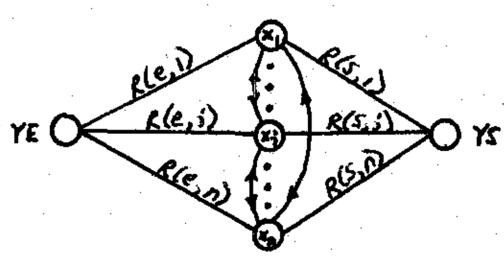
The lens model represented in Figure 1 will now be developed. Before getting started however, note that the distal variable side of the lens model may not be known. If the distal variable side is not known, the value of the formulation is less clear. Knowledge of YE is central to the statistics of the lens model. For example, the achievement of the clinician is defined as the simple correlation between YE and YS and is denoted RA. The formulation, therefore, is for cases where YE can be determined with some accuracy. In many decision making situations, however, YE is not known. Rather some error prone estimate of YE is the only feedback available. For example, feedback on a patient's survival time, after estimating disease severity as YE, is error prone. The effect of the amount of error in feedback is one of the major topics to be considered in this thesis and it will be analyzed later.

To start to develop the lens model a decision making situation will be considered. The clinician confronted with a decision making situation will consider his cues. Denote the cues as x_i and a schematic of the cues can be represented as follows:



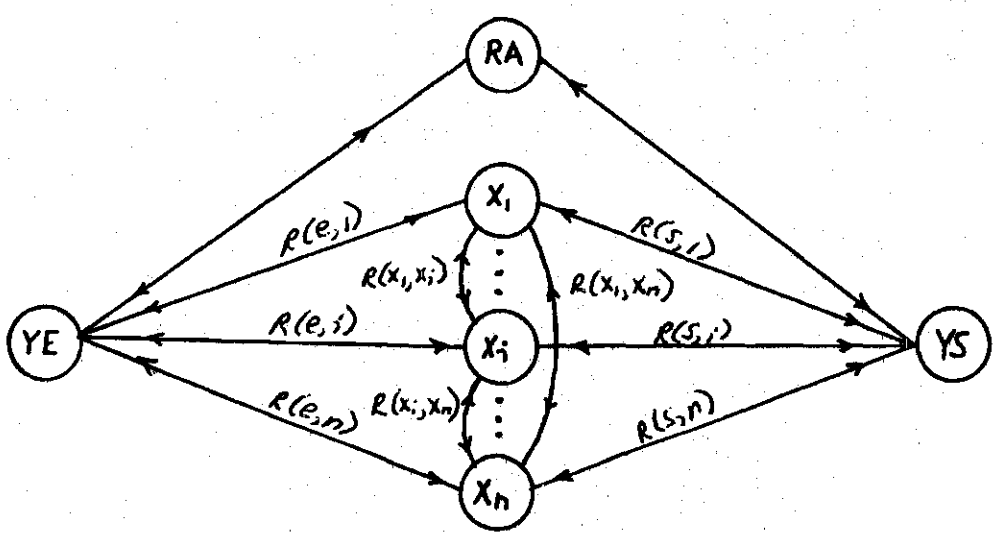
where $R(x_i,x_j)$ are intercorrelations

In considering the cues, the clinician will consider past and present information available to him. The clinician in considering his information will try to diagnose the patients present and future states. The clinician's decision is his estimate and is denoted YS. Because his decision may be wrong the actual state of the patient is called a distal variable and is denoted as YE. The schematic can now be expanded to the following:



where R(e,i) is called the ecological vality of the i^{th} cue and it is the simple correlation between YE and x_i. R(s,i) is called the utilization coefficient of the i^{th} cue and it is the simple correlation between YS and x_i.

Once a decision is made the clinician must wait to observe the YE that occurs. The correlation of YS to YE over time can be looked on as achievement and is denoted as RA. The final schematic below is called the lens model.



where x_i, YE, YS, R(e,i), R(s,i) and RA are as defined earlier.

Further Development of Lens Model Correlations

A mathematical formalization of the general lens model was provided by Hursch, Hammond and Hursch, (1964), and refined by Tucker, (1964). The major points of these formalizations are summarized in Appendix IV of the present study.

Of the various statistics available for describing the problem-response system represented by the lens model, three are of particular interest for the present research. The first achievement, denoted as RA is defined as the correlation between the distal variable and the subject's response. Achievement is the appropriate word for this correlation since it gives a simple over all measure of the subject's ability to estimate distal variable values in a given problem. A second measure G, is the simple correlation between $\hat{YE}$, the best fit linear prediction of YE from the cues, and $\hat{YS}$, the best fit linear prediction of YS from the cues. Therefore G, reflects whether or not the linearity of the subject's estimates of the distal variable matches the linearity that exists in the distal variable values. A linear model is used because it is known to be a good predictor in lens model type decision problems. Finally, the correlation between $\hat{YS}$ and YS indicates how well the linear model of the subject's estimates actually fits the subject's estimates. The correlation between $\hat{YS}$ and YS is denoted as RS.

The general lens model equation (Hursch, Hammond and Hursch, (1964)):

$$RA = G^{*}RS^{*}RE + C\sqrt{1 - RE^2}\sqrt{1 - RS^2}$$

shows that, for a highly linear problem, (RE approximately equal to one), high values of both RS and G are required for overall achievement to approach the upper limit RE. In a recent paper, Hammond and Summers, (1972), argue that G provides an index of the subject's "knowledge" of of the properties of the task, while RS provides an index of his "cognitive control" over that knowledge. Since G and RS may be shown to be statistically independent, this formulation offers the intriguing possibility of empirically disentangling the process of acquisition and application of understanding in complex inference tasks. Unfortunately, the formulation appears to be limited to linear tasks, for which

C is zero. Given the rather tentative status of the Hammond and Summers interpretation of G and RS their terminology will not be adopted here.

Some Illustrative Lens Model Research

The conceptual framework reviewed thus far has generated a very considerable body of empirical research. A recent partial review of this work (Slovic and Lichtenstein, (1971)) notes several hundred papers based on the lens model, and related work using Bayes Theorem as the central model. Clearly, no exhaustive review of this work is intended here; the interested reader is referred to such sources as Hammond, (1966), and Slovic and Lichtenstein, (1971). The present section will review a highly selected subset of this literature, with a view to suggesting the value of the lens model formulation as a stimulus to empirical research.

The first topic considered here will be the ability of people to learn to use linear cues appropriately in a multiple cue decision task. Lee and Tucker, (1962), studied this first topic by putting subjects in a situation where they would have to define a card handed them as being an X card or not an X card, where X cards had certain characteristics that differentiated them from non X cards. Three different conditions were studied. The results of Lee and Tucker, (1962), led them to the conclusion that people can combine pertinent cues and that the people learned to weigh the more important probabilistic cues more, in making judgements.

Uhl, (1963), studied a situation where subjects tried to predict a response resulting when three stimuli occurred. The subjects looking at rows of lights had to predict the correct criterion light that would come on. The criterion light was related to the three stimuli lights by a multiple regression equation. This meant the subject had to learn the beta weights. The results showed that subjects can learn to use the row of lights to predict the criterion light. It should also be noted that the subjects had no knowledge that the solution was of multiple-regression form.

Summers, (1962), presented subjects with different cues simultaneously in an effort to try to determine if responses came to depend on a particular cue. Summers found that subjects tended to learn to use correct cues. The subjects learned to develop weights for cues that were close to optimal weights for each cue.

These and similar studies such as Newton, (1965), Hammond, Hursch and Todd, (1964), and Todd and Hammond, (1965), suggest that people do have the ability to learn how to use linear cues. The studies indicate that subjects can develop appropriate weights. In other words, subjects learn which probabilistic cues are more important.

Another area of interest is the difference in learning ability in nonlinear and linear functional relationships. Summers, Summers and Karkau, (1969), studied this area by having the subjects deal with two cues. The two cues were related to the criterion variable by four different functional relationships. The four functional relationships were each defined as a condition requiring the prediction of "age" from two numerical cues, c_1 and c_2, and appear as the following:

$$\text{Condition 1;} \quad \text{age} = \frac{100}{14}(C_1 + C_2 - 2) + V$$

$$\text{Condition 2;} \quad \text{age} = \frac{100}{63}(C_1C_2 - 1) + V$$

$$\text{Condition 3;} \quad \text{age} = \frac{100}{7}(C_1^{C_2/8} - 1) + V$$

$$\text{Condition 4;} \quad \text{age} = \frac{100}{7}(C_2^{C_1/8} - 1) + V$$

The results of this study indicate that a difference in learning occurred for each of the conditions. However, performance increased for all types of feedback but the accuracy of the response over conditions differed. Subjects in nonlinear conditions learned to use the cues to predict the criterion variable. Thus this study indicates that subjects learn to predict the criterion variable no matter how the criterion variable is related to the cues.

Hammond and Summers, (1965), divided their subjects into three groups. The experimental procedure within the groups was as follows:

Group 1. "The subjects in group 1 were told simply to make inferences on the basis of the two test scores and that accurate prediction was possible." (page 219).

Group 2. "The subjects in group 2 were told that both linear and nonlinear relationships were involved and that the highest accuracy possible was contingent upon utilization of both." (page 219).

Group 3. "The subjects in group 3 were also provided with information regarding linear and nonlinear relationships; in addition, the linear and nonlinear cues were identified." (page 219).

The subjects were studied in the following task:

(a) " . . . one cue related in a linear, the other in a nonlinear manner to a criterion." (page 215).

(b) " . . . the criterion partly, but not perfectly, predictable from either cue alone." (page 215).

(c) " . . . the criterion perfectly predictable from appropriate utilization of both." (page 215).

The following quotes summarize their results:

> 1. . . . the nature of the information given the subject determines whether the linear or nonlinear cue will be increasingly utilized during learning. There is a much stronger dependence on the linear cue than on the nonlinear cue in the minimum information condition. (page 221).
>
> 2. Overall, there appears to be a tendency to depend more on the linear cue. (page 221).
>
> 3. Although results of previous studies indicate a high degree of linearity on the part of the subjects in multiple-cue probability learning tasks, the present results indicate that the propensity for a highly linear, additive response system is contingent upon the subject being presented with a highly linear task system. (page 222).

The last paper to be considered in this area is Summers and Hammond, (1966). This paper uses the same task as the Hammond and Summers, (1965), paper. The difference occurs in the purpose of the research. The purpose of Summers and Hammond, (1966), was to

> . . . study multiple-cue probability learning in tasks involving nonlinear as well as linear cue-criterion relations and to study the effects of different levels of linear and nonlinear task variance and different task instructions upon, (a) Subject's inductive achievement and (b) Subject's cue dependence in a task which requires the utilization of both linear and nonlinear cue-criterion relations if perfect accuracy is to be achieved. (page 753).

Summers and Hammond, (1966), report:

> The major findings of this study are that (a) both achievement and cue dependence are affected by task information and task properties and that (b) subjects can learn to make inferences from nonlinear as well as linear task relations. (Page 756).

Thus, while several studies have examined the impact on performance of outcome and cognitive feedback it appears that no available research has examined the effect of error prone outcome feedback in these tasks. Since it seems that feedback will be significantly error prone in a wide range of real-world tasks the effect of such feedback on learning and performance is of considerable interest. In the following section, two research questions and related hypotheses concerning these effects will be developed. A research design for experimental investigation of these hypotheses will be presented in the next chapter.

Research Questions

Brunswik, (1952), said feedback kept the organism from a state of isolation. In other words, an organism will use previous results to guard against the possibility of misinterpreting input signals. For example, consider a situation where a person is to combine two cues so that his method of combination results in a response that is close to the response of some model unknown to him. Then according to Brunswik, (1952), the subject will use his feedback to develop a better method for predicting this unknown model. If his feedback is error prone, the subject's model should decrease in accuracy as the error in feedback increases. If the amount of error in feedback varied, we would expect high misinformation to result in lower achievement than low misinformation. The first question to be asked is the following:

Q1: Does a change in misinformation cause the subject's achievement (RA) values to decrease?

To test a level of misinformation given to a subject the YE variance must be controlled. Misinformation will be denoted by L and it will be defined as the percentage of feedback variable variance not accounted for by the variance of YE. Let

YE = distal variable

YF = feedback variable

EF = error in feedback

Since YE is controlled its distribution, mean and variance are known. Also EF is controlled and the YE and EF values are generated independently of each other. Define YF = YE + EF, then from the independence of YE and EF, it is known that Var(YF) = Var(YE) + Var(EF). Independence enables the variance of YF to be separated into two parts. Since misinformation was defined as the percentage of feedback variable variance not accounted for by the variance of YE the L value can be defined as follows:

$$L = \frac{\mathrm{Var(EF)}}{\mathrm{Var(YF)}} \times 100\%$$

The value of L can thus be adjusted by merely changing Var(EF), the error variance included in the feedback. The achievement correlation values RA will be studied at different L values, for different blocks of trials.

In Appendix IV the RA value is shown (following Tucker, (1964)) to be equal to the following:

$$RA = G^{*}RS^{*}RE + C\sqrt{1 - RE^{2}}\sqrt{1 - RS^{2}} \tag{1}$$

where

G is the correlation between $\hat{YS}$ and $\hat{YE}$ which are estimates from their respective regression equations.

RE is the multiple correlation coefficient between YE and $\hat{YE}$.

RS is the multiple correlation coefficient between YS and $\hat{YS}$.

C is the correlation between the residuals of the regression equations yielding $\hat{YE}$ and $\hat{YS}$.

Since there is no systematic non-linearity in this task C will approach zero (except for sampling error), and equation (1) reduces to

$$RA = G^{*}RS^{*}RE \tag{2}$$

Hammond and Summers, (1965), define the linearity of subject's response as the correlation between $\hat{YS}$ and $\hat{YE}$. This means G can be defined as the linearity of a subject's response where response is used because G is an indication of how the regression equation of YS fits the regression equation YE.

A pertinent question then is the following:

Q2: What is the effect of misinformation on the subject's G and RS correlation values?

Hypotheses to be Investigated

Brunswik, (1952), stated that people tend to change states of behavior. This means that if a subject is getting error in feedback he may change his model to fit his incorrect feedback. The result of a subject changing his model to fit error in feedback might cause a decrease in the subject's achievement. This decrease would be due to error in feedback. The first hypothesis then will be as follows:

Hyp. 1: As misinformation increases in a multiple-cue probability learning task under error prone outcome feedback the achievement correlation will decrease.

From equation two it can be seen that if RA is high G and RS must also be high but if RA is low there are more possible values for G and RS. It would be interesting then to see whether RA varies with G or with RS. Hypothesis two is as follows:

Hyp. 2: In a multiple-cue probability learning task, the change in achievement over various levels of error in feedback will be primarily due to the failure of the subject's estimates to fit the linear regression equation of his estimates, than to the failure of the linear regression equation estimates of the subject's predictions of the distal variable to correlate with the linear regression equation estimates of the distal variable values.

Another way of stating hypothesis two is to say that the change in RA will be primarily due to RS than to G. The next chapter will deal with the experimental design.

CHAPTER II

METHODS AND PROCEDURES

Problem

The subjects were to make decisions based on the information they were provided. They saw three numbers representing point spreads of football games that three individual experts published. The subjects were told that the experts made their decisions based on perfect game conditions. (Where perfect game conditions implies beautiful weather, no injuries and good officials, etc.). Also each subject always knew which prediction went with which expert.

The subjects, therefore, saw the three expert predictions and then they made their own prediction. Once the subject's prediction was recorded feedback was given to the subject. The point spread of the game along with the conditions under which the point spread occurred were told to the subject. No names were attached to any teams in an effort to eliminate bias. The conditions mentioned above that a subject received were used to introduce error in feedback. For example, three types of game conditions were defined for the subjects as follows:

Normal Game

No unusual factors affect the result. This means weather conditions are good, injuries are few, etc. Therefore, we would expect the experts pre-game predictions to be reasonably good.

Somewhat Unpredictable Game

A number of unusual incidents occurred such as key injuries, poor weather conditions, etc. These unusual incidents influenced the outcome of the game somewhat. We, therefore, expect the outcome of the game to be somewhat different from the expert's pre-game predictions.

Very Unpredictable Game

Many unusual incidents effect the outcome of the game. Therefore, the outcome of the game may be considerably different from the experts pre-game predictions.

The following example was shown to the subjects:

Example. Suppose I see the following:

Expert 1	10
Expert 2	5
Expert 3	8

Considering the above I predict the point spread will be 6 which means I fill in the answer form as follows:

#	Your Prediction	Actual Result
1	6	
2		

After writing in the 6, I am told it was a Very Unpredictable Game and the point spread turned out to be 20. Then for this experimental run my data sheet looks as follows:

#	Your Prediction	Actual Result
1	6	20
2		

I am now ready to look at three more expert values.

Experimental Design

The distal variables generated represented point spreads of football games. The point spreads were drawn from a normal distribution with mean zero and variance two-hundred and twenty-five, so that they were between a plus or minus fifteen points approximately sixty-seven percent of the time.

It was decided that the subjects would see three numbers representing three different expert opinions as to the actual point spread that would occur for a football game. These three cue values representing expert judgements of the point spreads were generated as follows:

$$X_1 = YE \pm E_1$$

$$X_2 = YE \pm E_2$$

$$X_3 = YE \pm E_3$$

Where X_1, X_2, and X_3 represent experts and E_1, E_2, and E_3 represent error. In a previous chapter it was mentioned that the YE and E_i are independent.

The subjects were divided into two different groups. The first group saw all N(0, 26) experts while the second group saw two N(0, 36) experts and one N(0, 100) expert. A N(0, 36) expert implies the error term E_i comes from a normal distribution with mean zero and variance thirty-six. Because of independence, the variance of the experts predictions is Var YE + Var E_i, which means for higher variance in E_i the corresponding expert X_i should be weighed less. In other words, the independence enables us to see just how well the second group was able to deal with a situation with less reliable information to use.

It should be noted that the X_i are linear functions of YE with a small error term. This type of relationship between X_i and YE was noted in Chapter I to result in a simililified equation for RA. Therefore, the relationship between YE and the X_i will enable us to study the hypotheses.

When subjects arrive at a response they can do so using either an analytical or an intuitive model. The analytical model is one where the subject is allowed to develop some formulation such that when given inputs for his formulation he can use external sources such as paper, pencil, calculators or other computers, while the intuitive model does not allow the subject to use these external sources. An intuitive model was imposed on the subjects because if a subject used an analytical model he may not change his model as his intuition tells him to change. Thus to study the correlation between $\hat{YE}$ and $\hat{YS}$, the correlation between $\hat{YS}$ and YS and achievement an intuitive model was imposed on the subjects. The intuitive model enables us to study the latter points by the fact that subjects will change an intuitive model at higher levels of error where if the subjects used an analytic model they may not change the model.

After a subject recorded his estimate, he was given feedback. The levels of error in feedback were zero, twenty-five and fifty percent. The levels mentioned were chosen because it was felt if a trend existed those levels would indicate the trend. It was also felt that a subject starting with zero error and proceeding to fifty percent error may react differently than a subject starting at fifty and proceeding downward. The trend that may occur in going from fifty to zero can be checked by taking the subject back to fifty percent error. So for the two groups we had the following design:

Experiment 1 (3 experts N(0, 36))

Group A 0 25 50 25 0

Group B 50 25 0 25 50

Experiment 2 (2 experts N(0, 36); 1 expert N(0, 100))

Group Ax 0 25 50 25 0

Group Bx 50 25 0 25 50

An example of just how the error in feedback is determined will be given for twenty-five percent. It should be noted the other levels are calculated similiarly. The twenty-five percent error is arrived at as follows:

a.) Call up a random number, say Q1.

b.) If Q1 is less than or equal to 0.5 set Q1 equal to a totally new positive random number. If Q1 is greater than 0.5 set Q1 equal to a totally new negative random number. (This gives fifty percent plus or minus and the values are equally likely between plus and minus one.)

c.) Now let X represent the variance of the error value. Then for this case

$$.25 = \frac{X}{X + \mathrm{Var(YE)}} \qquad .75X = (.25)^{*}(\mathrm{Var(YE)})$$

But the Var(YE) = 225. So

$$X = \frac{56.25}{.75} = 75.000$$

or the standard deviation of X = 8.635.

d.) The feedback value of YF is then set equal to the following:

$$YF = YE + Q1^{*}(8.635)$$

The experimental design is completed and the method of experimentation follows.

Method of Experimentation

Subjects

The subjects were undergraduates at Georgia Institute of Technology. They were tested in two main groups. Group one consisted of fifteen males and one female. Group two considted of nineteen males and one female. The groups were structured as follows:

Experiment I (All 3 experts N(0, 36))

A	B
8 subjects (0 25 50 250)	8 subjects (50 25 0 25 50)

Experiment II (2 experts N(0, 36); 1 expert N(0, 100))

Ax	Bx
10 subjects (0 25 50 25 0)	10 subjects (50 25 0 25 50)

Apparatus

The apparatus was a cardboard box illustrated below.

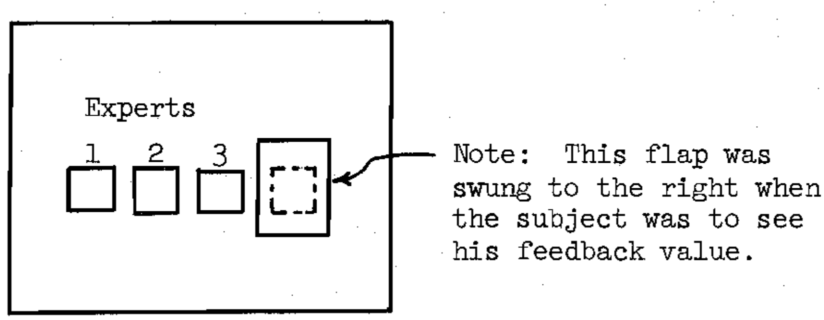


Figure 2. Front View of Apparatus

Inside the box was placed a roll of paper which contained the cue values written in black and the feedback value written in red. There were six rolls of data prepared, three for experiment one and three different rolls for experiment two. Each roll represented thirty trials at a particular level of error. The roll was placed in the box and the beginning of the roll was pulled upward till three new cue values appeared. The procedure was fast and worked quite well. A roll appears as follows:

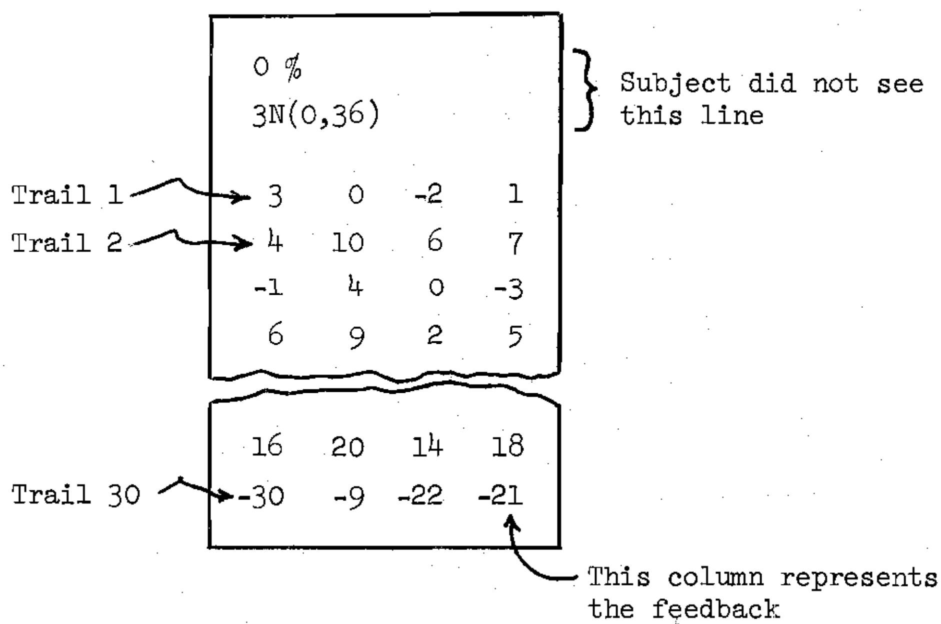


Figure 3. Data Sheet

Instructions

The first group tested was group one. A set of instructions was handed to each individual in the group. The actual instructions are included in Appendix I. The instructions were read to the group as a whole. It should be noted that as the instructions were read the experimenter pointed to the apparatus where the subject would see the numbers about which he was reading. After reading the instructions, any questions were answered and pencils and answer forms were handed out. At this point the group was divided into set A and set B. This division was accomplished by merely placing the subjects in the left side of the room in set A and placing those not in set A in set B.

Procedure

Once the subjects understood the instructions group B was asked to go th another room. The experiment was then started on group A. Group A first saw thirty trails at zero percent error. For each trail each expert value was called out from left to right. The cardboard flap was swung to the right after all the subjects had written down their estimates. As the flap was swung the experimenter stated team A won or lost by the points indicated. The experimenter also indicated at this time the conditions under which the outcome occurred. Note here that a condition is used as a means of conveying the level of error present. We did this because it was felt the subject would have a better feel of the error present. Also note team A was not associated to any team because we wanted no bias to enter any given decision made. After thirty trials at zero percent group A proceeded to twenty-five and then to fifty percent error. At this time, group B entered the room and thirty trials at fifty percent were run on both groups. The complete group also had twenty-five and zero percent error run on them, in that order. After the zero percent error, group A was asked to leave and the experiment was completed on group B at twenty-five and fifty percent error. Note when group B re-entered the room the instructions were reviewed quickly and at all times the subjects could see the numbers called out. It might also be of interest to know the subjects were kept occupied when they were not involved directly with the experiment. An illustration of the above is below.

```
                         A Leaves
A      0 25 50 25 0                A
B            50 25 0 25 50         B
  B Leaves
           B Enters
```

Note each 0, 25 or 50 percent error represents a roll of numbers so each subject saw the same sets of numbers.

The experiment on group two was performed exactly like group one.

CHAPTER III

RESULTS

Introduction

This chapter is divided into two sections. The first section will deal with testing the hypotheses while the second section will deal with other results that the data yielded. Note that the calculations of Section I are found in Appendix III in detail along with the data.

Section I

Hypothesis 1 was previously stated as follows:

Hyp. 1: As misinformation increases in a multiple-cue probability learning task under error-prone outcome feedback the achievement correlation will decrease.

The data in groups A and B were pooled because at each level of misinformation both groups saw the same values, meaning there was homogeneity of variance at each level. R. A. Fisher's Z-transformation was applied to the raw correlation coefficients before ANOVA was carried out. The data considered here are transformed achievement correlations obtained over levels of misinformation. A one-way ANOVA was used to test if there was a difference in achievement over the levels of misinformation. The results appear in Table 1.

The F-ratio is significant at $\alpha = 0.05$. This means a difference in achievement did occur for different levels of misinformation. A

Table 1. Mean RA by Level of Misinformation for Pooled Groups A and B

Level of Misinformation	Mean RA Values	Sample Size	F-Ratio
0%	0.963	24	3.3149
25%	0.960	32	
50%	0.954	24	$P \leq 0.05$

Duncan Multiple Range (DMR) test at $\alpha = 0.05$ indicated that:

0% different from 25% and 50%

25% and 50% were not different

Where 0% implies error free feedback and 25% and 50% imply that the feedback was error prone.

The above results support Hypothesis 1. For groups Ax and Bx, where the "bad" expert was added the trends are much more distinct. The results for the pooled groups Ax and Bx appears in Table 2.

Table 2. Mean RA by Level of Misinformation for Pooled Groups Ax and Bx

Level of Misinformation	Mean RA Values	Sample Size	F-Ratio
0%	0.956	30	22.9734
25%	0.897	40	
50%	0.798	30	$P \leq 0.01$

The F-ratio is significant at $\alpha = 0.01$. This means a difference in achievement did occur for different levels of misinformation. A DMR test at $\alpha = 0.05$ indicated that:

0% different from 25% and 50%

25% different from 50%

This means achievement differed over all levels of misinformation as predicted by Hypothesis 1.

Therefore, from the analysis of the above groups Hypothesis 1 is strongly supported by these data.

The second hypothesis was previously stated as follows:

Hyp. 2: In a multiple-cue probability learning task, the change in achievement over various levels of error in feedback will be primarily due to the failure of the subject's estimates to fit the linear regression equation of his estimates, than to the failure of the linear regression equation estimates of the subject's predictions of the distal variable to correlate with the linear regression equation estimates of the distal variable values.

To test this hypothesis the following correlations were calculated:

R(RA, G) - The correlation between achievement (RA) and the correlation of the linear regression equation estimates of the subject's estimates of the distal variable and the linear regression equation estimates of the distal variable.

R(RA, RS) - The correlation between achievement (RA) and the correlation of the linear regression equation estimates of the subject's estimates of the distal variable and the actual estimates of the distal variable by the subject.

The following values were obtained for the data for groups A and B:

$$R(RA, G) = 0.659436367 \quad \text{denote } R(RA, G) \text{ as } \hat{\rho}_1$$

$$R(RA, RS) = 0.946386516 \quad \text{denote } R(RA, RS) \text{ as } \hat{\rho}_2$$

Notice that R(RA, RS) > R(RA, G). To test the statistical significance of this difference we apply a Z-transformation and use the data listed in Table 3.

Table 3. Data Needed to Calculate W

	1	2
$\hat{\rho}_i$	0.659436367	0.946386516
z_i	0.7918156483	1.79596463
$n_i - 3$	77	77
$z_i(n_i - 3)$	60.96980492	138.2892765

$$W = \sum_{i=1}^{2} (n_i - 3)(z_i - \bar{z})^2 \text{ distributed as } X^2 \text{ 1 df}$$

$$W = 38.7756059$$

W can not be accepted at $\alpha = 0.0001$. This means that $R(RA, G) \neq R(RA, RS)$. Therefore, hypothesis two is verified for groups A and B.

For groups Ax and Bx the following correlations were obtained:

$$R(RA, G) = 0.705461256 \text{ denote } R(RA, G) \text{ as } \hat{\rho}_1$$

$$R(RA, RS) = 0.985904209 \text{ denote } R(RA, RS) \text{ as } \hat{\rho}_2$$

As before R(RA, RS) > R(RA, G). Again, a Z-transformation was applied to test for a significant difference between R(RA, G) and R(RA, RS). The data needed to test H_o; $\rho_1 = \rho_2$ can be found in Table 4.

Table 4. Data Needed to Calculate W

	1	2
$\hat{\rho}_i$	0.705461256	0.984904209
z_i	0.8780901658	2.473976684
$n_i - 3$	97	97
$z_i(n_i - 3)$	85,17474608	239.9757383

$$W = \sum_{i=1}^{2} (n_i - 3)(z_i - \bar{z})^2 \text{ distributed as } X^2 \text{ 1 df}$$

$$W = 123.5194205$$

W can not be accepted at $\alpha = 0.0001$. This means that R(RA, G) $\neq$ R(RA, RS). Therefore, hypothesis two is verified for groups Ax and Bx.

There is a significant difference between R(RA, G) and R(RA, RS) in both groups A and B and Ax and Bx. Also for both sets of groups, the correlation between achievement and the correlation of the linear regression estimates of the subject's estimates of the distal variable and the actual estimates of the distal variable by the subject (R(RA, RS)) was higher indicating that performance degradation is primarily due to the failure of the subject's estimates to fit the linear regression equation of his estimates, than to the failure of the linear regression equation estimates of the subject's predictions of the distal variable to correlate with the linear regression equation estimates of the distal variable values. More briefly, using the Hammond and Summers (1972) terminology discussed earlier, we could say that, in this task, performance degradation with increasing feedback error is primarily due to loss of "cognitive control" (RS) rather than to loss of "Knowledge" (G).

In this section both major hypotheses have been supported. Secondary analyses of the data will be presented in the following section.

SECTION II

Group A

It was previously stated that group A subjects were presented misinformation in the following order: zero, low, high, low, zero. The correlations RA, G, and RS where totaled and averaged. The averages appear in Table 5 and are plotted in Figure 4.

The value of G stays very stable over all levels of misinformation. Since G is very close to one, the subject's estimates must be approximately as linear as the distal variable values. However, RS is less than G indicating that the subject's estimates are not completely linear. Note RS is relatively constant as is RA. The effect of misinformation in feedback on group A can not be clearly seen.

Table 5. Average Values Over Subjects of Group A

	0	25	50	25	0
RA	.970	.956	.950	.954	.946
G	.997	.996	.992	.996	.992
RS	.984	.981	.976	.983	.965

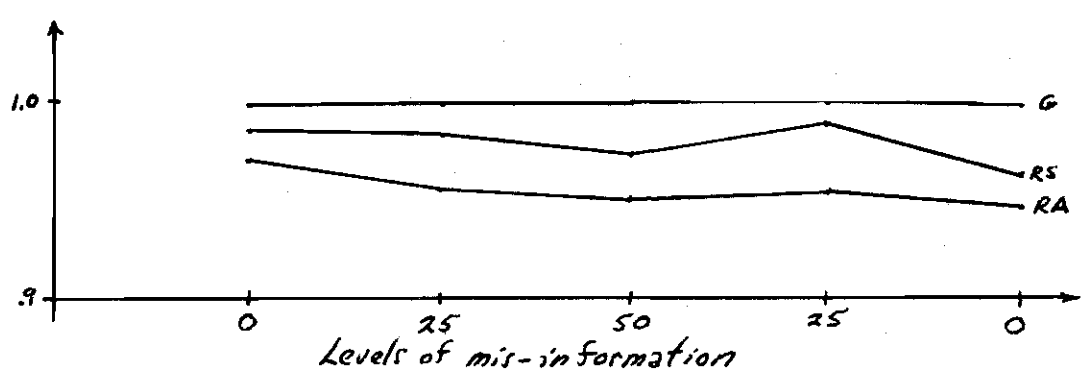


Figure 4. Average Values G, RS, and RA of Subjects of Group A

<u>Group B</u>

Subjects in group B were presented misinformation in the following order high, low, zero, low, high. The RA, RS, and G values were averaged over all the subjects of group B. The averages appear in Table 6 and are plotted in Figure 5.

Table 6. Averages Over Subjects of Group B

	50	25	0	25	50
RA	.946	.961	.975	.965	.965
G	.995	.998	.997	.998	.995
RS	.967	.986	.995	.985	.984

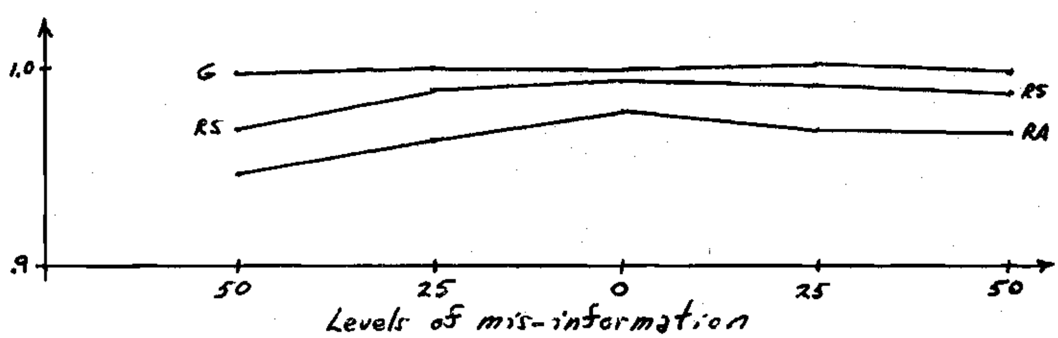


Figure 5. Average Values G, RS, and RA of Subjects of Group B

RS approaches G very fast, however, RS decreases slightly as the level of misinformation increases. The effect of misinformation in feedback on group B cannot be clearly seen.

Groups A and B do show slight trends that indicate misinformation hinders performance. Because of the one "bad" expert in groups Ax and Bx the slight trends should become distinct trends. As will be seen this is exactly what happened indicating that the added source of error hindered the subjects.

Group Ax

The difference between group Ax and Group A was that group Ax had one "bad" expert. This "bad" expert had a larger error variance than the other experts. The correlations RS, RA, and G were averaged overall the subjects of group Ax. The averages appear in Table 7 and are plotted in Figure 6.

Table 7. Average Values Over Subjects of Group Ax

	0	25	50	25	0
RA	.958	.912	.885	.926	.962
G	.996	.989	.979	.988	.998
RS	.978	.961	.923	.964	.979

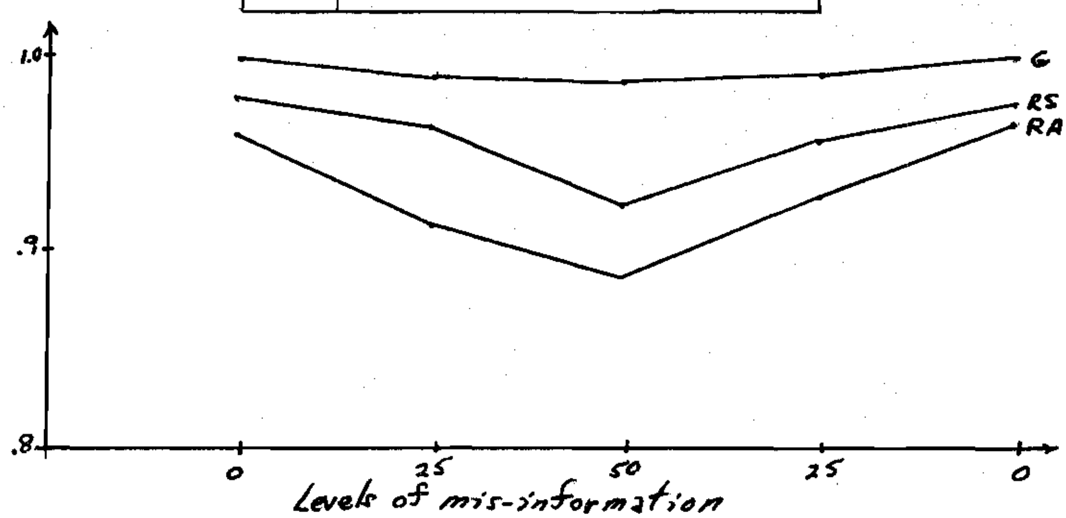


Figure 6. Average Values G, RS, and RA of Subjects of Group Ax

The level of misinformation here has a definite effect on RA, RS, and G. It is noted that as the level of misinformation increases RA decreases and as the level decreases RA increases. The correlation between the linear regression equation estimates of the subject's predictions of the distal variable and the linear regression equation estimates of the distal variable values seems to remain fairly constant but the correlation between the subject's estimates and the linear regression equation of his estimates seems to vary inversely with level of misinformation.

Group Bx

Group Bx had the same added source of error as Group Ax. The correlations RS, RA, and G were averaged overall the subjects of group Bx. The averages appear in Table 8 and are plotted in Figure 7.

Table 8. Average Values Over Subjects of Group Bx

	50	25	0	25	50
RA	.795	.897	.949	.852	.715
G	.952	.983	.991	.990	.961
RS	.854	.937	.973	.902	.792

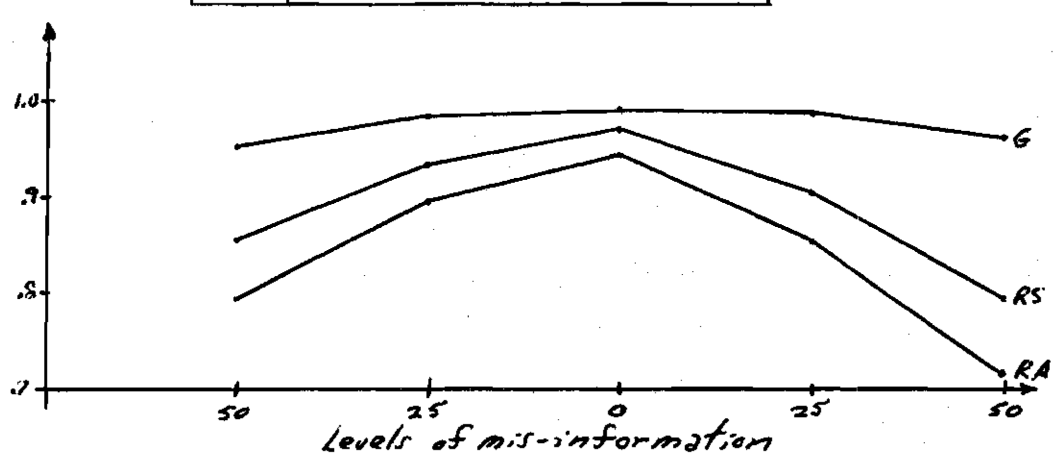


Figure 7. Average Values G, RS and RA of Subjects of Group Bx

RS and RA are seen to vary inversely with misinformation. However, it is noted that the trends are much more distinct for Group Bx than group Ax. This may have occurred because the subjects of group Bx starting at a high misinformation level never developed an intuitive model of the problem. At any rate, misinformation clearly hindered achievement and the correlation between the subject's estimates and the linear regression equation of his estimates.

CHAPTER IV

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Subjects With Three N(0,36) Experts

In section II Figures 4 and 5 indicate that subjects realized a linear model was best. However, the RS values of the subjects are generally below their G values. This means the subject was using a model that did not predict the distal variable as well as a linear model would have. The following major hypotheses were supported:

Hyp. 1: As misinformation increases in a multiple-cue probability learning task under error prone outcome feedback the achievement correlation will decrease.

Hyp. 2: In a multiple-cue probability learning task, the change in achievement over various levels of error in feedback will be primarily due to the failure of the subject's estimates to fit the linear regression equation estimates of the subject's predictions of the distal variable to correlate with the linear regression equation estimates of the distal variable values.

Subjects With Two N(0, 36) Experts and One N(0, 100) Expert

The one "bad" expert added another source of error to the problem. Because of this added error clear effects of misinformation were expected. As can be seen in Figures 6 and 7 of Section II this is precisely what happened.

As the subjects went from zero to high misinformation his G and RS values dropped. The subject's linearity of prediction seemed to improve rapidly as the level of misinformation decreased. It was clear that misinformation did cause the subject's predictions to decrease in accuracy.

For subjects starting with high misinformation G starts low and increases rather fast. However, the second time the subject sees increased misinformation his G values seem to decrease less rapidly. But the subject's linearity of prediction decreases faster. This is probably caused by the fact the subject's did not develop a linear model. The effect of misinformation again was to hamper the subject.

The level of misinformation definitely affects the subject's achievement. Subjects tend to achieve less at high misinformation levels than at low levels. This decrease in achievement was due primarily to the failure of the subject's estimates to fit the linear regression equation of his estimates. A subject, therefore, seems to use a non linear model when he receives error in feedback. The subject's seem to know a linear model predicts best but they just cannot apply it correctly.

The subjects in groups Ax and Bx tended to fluctuate much more than groups A and B. Fluctuate is meant to imply that trends are more distinct. This implies that subjects beginning at low levels of error first tend to develop a model that they can apply at higher levels of error.

Summary of Findings

The data reported here provided strong support for the two major hypotheses: that performance is degraded by increasing feedback error and this degradation is primarily attributable to the failure of the subject's estimates to fit the linear regression equation of his estimates under high error in feedback conditions. The impact of error in feedback on the correlation between the linear regression equation estimates of the subjects predictions of the distal variable and the linear regression equation estimates of the distal variable values was hardly noticeable.

Two suggestive secondary findings should also be noted. First, there is some evidence that performance degradation with increasing feedback error is smaller for subjects who initially learn under error-prone feedback (Groups A and Ax, Blocks 1-3) than for those who initially learn under error-prone feedback (Groups B and Bx, Blocks 3-5). Second, there is a strong suggestion in these data that the phenomena noted here are highly sensitive to changes in problem structure. The apparently minor change of introducing slightly more error in one cue led to considerably more variation over treatments in all three major measures of performance in Experiment 2 than in Experiment 1.

Recommendations

The findings of this study can be applied to several areas. Two of these areas will be discussed. First, consider the interface of man and machine. Goldberg, (1962), Bowman, (1963), and Kunreuther, (1969), indicate that machines can be programmed with models that out-perform man in certain decision making situations. In this study, it was found that a linear model outperformed the subject. However, the linear regression estimates of the subject's predictions were highly correlated with the linear regression estimates of the distal variable. This means the subject's estimates should have been very good predictors, but the subject's estimates were not that accurate. A possible reason for the subject's failing to use only a linear model was given by Goldberg, (1968), and Bowman, (1963). Bowman, (1963), reports that

> . . . Man seems to respond to selective cues in his environment - particular things seem to catch his attention at times (the last telephone call), while at other times it is a different set of stimuli. Not only is this selective cueing the case, but a threshold concept seems to apply. He may respond not at all up to some point and then overrespond beyond that. It is this type of behavior which helps explain the variance in the organization's (or its management's) behavior. (page 316).

Bowman, (1963), clearly points out that man has inconsistencies which are not found in machines. The question to consider is how man can use the machine to better his decision making model. Of course, man does not know the model that he uses but possibly practice making decisions with a lens model type program will have a lasting effect on the subject. The persistence over time of such improved performance is an area for further empirical study.

The second area to consider is the type of informational milieu in which the manager should be trained. It was seen that subjects starting with high error in feedback did not do as well when returned to the error condition as those starting with zero error in feedback. These data would indicate that subjects should be trained with the "best" feedback possible. The amount of feedback that is needed should be studied. Also it would be interesting to see if a subject who practiced at zero error in feedback over five blocks would outperform a subject who practiced at a low level of error in feedback when both subjects were tested over several blocks at a high level of error in feedback. The present data provides some tentative support for training decision makers in "classroom" milieus where error-less feedback can be provided, rather than in "real-world" settings. Again further empirical study across a range of problems, settings, and subjects would be required before this suggestion could be applied practically with much confidence.

APPENDIX I

INSTRUCTIONS

You are about to participate in an experiment in which you will make decisions based on the information you are provided. You will see three numbers. These numbers will represent point spreads of football games that three individual experts have published. Now the experts made their decisions based on perfect game conditions. (Where perfect game conditions implied beautiful weather, no injuries and good officals, etc.). You'll be told which number goes with which expert.

You are to write down your prediction of the point spread based on perfect game conditions. Once you have made an estimate, you'll be told the actual result of the game along with the conditions under which the point spread occurred. For example, a quarterback may have broken his collarbone in the first quarter. Because of this incident the expert's estimates along with your may be different from the point spread that resulted. However, you made your estimate without knowing, the quarterback would break his collarbone so don't get discouraged if you differ from the point spread that occurred.

You'll receive answer forms and pencils. Please note you are not allowed to write anything down except your estimate of the point spread and the feedback value you are told after you made your estimate.

#	Your Prediction	Actual Result
1		
2		

Before considering an example note that each time you see three expert estimates they represent estimates for different games. Now even the experts don't call every game exactly, but their predictions are sometimes better than others. We will consider three different kinds of games.

Normal Game

No unusual factors affect the result. This means weather conditions are good, injuries are few, etc. Therefore, we would expect the experts pre-game predictions to be reasonably good.

Somewhat Unpredictable Game

A number of unusual incidents occurred such as key injuries, poor decisions, poor weather conditions, etc. These unusual incidents influenced the outcome of the game somewhat. We therefore expect the outcome of the game to be somewhat different from the expert's pre-game predictions.

Very Unpredictable Game

Many unusual incidents effect the outcome of the game. Therefore, the outcome of the game may be considerably different from the experts pre-game predictions.

Example. Suppose I see the following:

Expert 1	10
Expert 2	5
Expert 3	8

Considering the above I predict the point spread will be (6) which means I fill in the answer form as follows:

#	Your Prediction	Actual Result
1	6	

After writing in the (6) I am told it was a Very Unpredictable Game and the point spread turned out to be (20). Then for this experimental run my data sheet looks as follows:

#	Your Prediction	Actual Result
1	6	20
2		

I am now ready to look at three more expert values.

APPENDIX II

DATA

Values Subjects Were Given

$N(0,36)$ for all experts

* Indicates subjects did not see that column

X_1	X_2	X_3	0%	X_1	X_2	X_3	25%	*0%	X_1	X_2	X_3	50%	*0%
0	-5	1	-5	0	-6	5	9	4	-20	-16	-14	-19	-13
18	7	7	5	-12	-12	-25	-18	-22	-13	-2	-19	7	-7
6	5	9	9	-25	-12	-21	-21	-19	-5	-7	-13	7	-4
11	18	-1	10	-22	-13	-9	-15	-13	-17	-21	-12	-27	-16
-20	-17	-21	-20	45	40	46	47	45	-16	-17	-23	-29	-20
34	24	27	31	17	12	14	7	9	6	-4	8	-11	1
-1	10	7	5	-7	-13	-14	-11	-13	31	27	26	37	26
11	6	14	7	1	-3	-1	8	1	-10	-15	-7	0	-11
18	25	26	26	0	-4	-5	-10	-2	-14	-5	-12	-25	-11
37	23	13	20	7	-10	-12	-2	-9	3	-1	2	13	4
-3	-1	1	-4	6	5	10	3	7	2	14	6	2	11
-25	-33	-36	-30	-9	-1	-7	6	2	-12	-24	-15	-5	-18
-4	-10	-10	-8	-3	11	6	-2	1	25	12	-1	9	13
3	13	3	8	40	28	34	24	33	13	-16	4	6	0
15	14	5	11	-1	-3	3	4	-2	-22	-12	-9	-13	-18
-5	-17	-19	-17	1	9	3	5	2	0	2	2	19	6
7	5	8	10	-15	-3	-17	-6	-9	-15	-25	-15	-27	-16
11	12	2	3	4	3	-4	3	0	-9	3	-11	-9	-7
1	-1	6	0	-1	-6	0	6	2	1	1	-2	-2	-1
8	6	14	12	-1	0	9	-4	4	4	3	-10	13	0
6	0	11	9	13	7	-2	9	10	-14	-22	-3	-23	-12
-12	-6	-20	-15	12	23	17	13	20	-5	-16	-14	-24	-10
-19	-8	-21	-17	-6	-1	-17	-13	-8	11	0	3	12	8
-25	-30	-16	-26	-13	-8	-13	-16	-11	31	20	13	5	16
-14	-9	-6	-5	-8	-5	2	-10	-2	22	34	36	26	32
28	45	39	37	3	6	2	8	2	-14	-2	-12	1	-9
25	17	15	17	-7	1	3	0	-5	5	8	0	13	3
-2	-3	3	1	-15	-15	-9	-14	-11	-2	-10	-9	-7	-10
11	10	6	10	-16	-13	-10	-15	-9	21	8	4	11	13
-6	-2	2	-8	2	5	-4	-3	1	-12	-16	-16	-20	-15

<u>Values Subjects Were Given</u>

N(0,36) for experts one and two ; N(0,100) for expert three

* Indicates subjects did not see that column

x_1	x_2	x_3	0%	x_1	x_2	x_3	25%	*0%	x_1	x_2	x_3	50%	*0%
0	-5	1	-5	0	-6	5	9	4	-20	-16	-14	-19	-13
18	7	8	5	-12	-12	-28	-18	-22	-13	-2	-26	7	-7
6	5	9	9	-25	-12	-23	-21	-19	-5	-7	-20	7	-4
11	18	-8	10	-22	-13	-6	-15	-13	-17	-21	-10	-27	-16
-20	-17	-23	-20	45	40	47	47	45	-16	-17	-25	-29	-20
34	24	25	31	17	12	17	7	9	6	-4	12	-11	1
-1	10	8	5	-7	-13	-15	-11	-13	31	27	25	39	26
11	6	18	7	1	-3	-2	8	1	-10	-15	-4	0	-11
18	25	26	26	0	-4	-7	-10	-2	-14	-5	-12	-25	-11
37	23	8	20	-7	-10	-13	-2	-9	3	-1	1	13	4
-3	-1	4	-4	6	5	12	3	7	2	14	3	2	11
-25	-33	-40	-30	-9	-1	-13	6	2	-12	-24	-13	-5	-18
-4	-10	-11	-8	-3	11	9	-2	1	25	12	-10	9	13
3	13	0	8	40	28	35	24	33	13	-16	6	6	0
15	14	1	11	-1	-3	7	4	-2	-22	-12	-3	-13	-18
-5	-17	-21	-17	1	9	4	5	2	0	2	-1	19	6
7	5	6	10	-15	-3	-23	-6	-9	-15	-25	-14	-27	-16
11	12	2	3	4	3	-7	3	0	-9	3	-13	-9	-7
1	-1	10	0	-1	-6	-1	6	2	1	1	-3	-2	-1
8	6	15	12	-1	0	13	-4	4	4	3	-16	13	0
6	0	12	9	13	7	-10	9	10	-14	-22	4	-23	-12
-12	-6	-23	-15	12	23	15	13	20	-5	-16	-17	-24	-10
-19	-18	-23	-17	-6	-1	-23	-13	-8	11	0	-1	12	8
-25	-30	-8	-26	-13	-8	-15	16	-11	31	20	11	5	16
-14	-9	-6	-5	-8	-5	4	-10	-2	22	34	39	26	32
28	45	41	37	3	6	2	8	2	-14	-2	-15	1	-9
25	17	14	17	-7	1	8	0	-5	5	8	-2	13	3
-2	-3	5	1	-15	-15	-8	-14	-11	-2	-10	-10	-7	-10
11	10	3	10	-16	-13	-10	-15	-9	21	8	-2	11	13
-6	-2	9	-8	2	5	-7	-3	1	-12	-16	-16	-20	-15

Correlations of the Data Given the Subjects

$N(0,36)$ For All Experts				$N(0,36)$ For Experts 1 and 2 $N(0,100)$ For Expert 3			
Ecological Validities							
$R(e,i)$	0%	25%	50%	$R(e,i)$	0%	25%	50%
$i=1$	0.9269	0.9405	0.9124	$i=1$	0.9269	0.9405	0.9124
$i=2$	0.9524	0.9290	0.9165	$i=2$	0.9524	0.9290	0.9165
$i=3$	0.9526	0.9467	0.9026	$i=3$	0.8755	0.8743	0.7581
Intercorrelations							
$R(x_i x_j)$	0%	25%	50%	$R(x_i x_j)$	0%	25%	50%
$x_1\ x_2$	0.8950	0.9016	0.7847	$x_1\ x_2$	0.8950	0.9016	0.7847
$x_1\ x_3$	0.8582	0.8862	0.8118	$x_1\ x_3$	0.7720	0.8156	0.6651
$x_2\ x_3$	0.8851	0.8758	0.7893	$x_2\ x_3$	0.7989	0.8072	0.6320
Betas							
β	0%	25%	50%	β	0%	25%	50%
1	0.2338	0.3341	0.3701	1	0.3018	0.4404	0.4432
2	0.3576	0.2458	0.3898	2	0.4685	0.3276	0.4583
3	0.4343	0.4366	0.2971	3	0.2663	0.2537	0.1718

Data Analysis for Group A

Subject 46	0	25	50	25	0	Subject 50	0	25	50	25	0
RS	.987	.988	.975	.983	.982	RS	.982	.982	.980	.967	.985
G	.994	.995	.994	.999	.995	G	.995	.996	.998	.991	.995
RA	.960	.969	.931	.950	.971	RA	.960	.973	.965	.992	.962

Subject 47	0	25	50	25	0	Subject 51	0	25	50	25	0
RS	.993	.965	.970	.978	.844	RS	.982	.978	.969	.969	.976
G	.998	.994	.999	.999	.965	G	.996	.999	.990	.998	.995
RA	.987	.945	.941	.955	.789	RA	.972	.962	.949	.944	.958

Subject 48	0	25	50	25	0	Subject 52	0	25	50	25	0
RS	.989	.991	.984	.988	.993	RS	.967	.981	.944	.982	.983
G	.999	.996	.993	.998	.997	G	.997	.996	.998	.999	.998
RA	.982	.961	.967	.967	.979	RA	.953	.949	.924	.946	.980

Subject 49	0	25	50	25	0	Subject 53	0	25	50	25	0
RS	.989	.984	.989	.996	.992	RS	.985	.979	.993	.999	.998
G	.999	.992	.968	.983	.993	G	.998	.999	.999	.999	.998
RA	.975	.942	.944	.954	.998	RA	.968	.950	.978	.977	.980

Data Analysis for Group B

Subject 31	50	25	0	25	50
RS	.964	.992	.991	.998	.984
G	.996	.998	.945	.999	.995
RA	.950	.966	.973	.974	.961

Subject 35	50	25	0	25	50
RS	.996	.997	.998	.992	.993
G	.999	.999	.997	.996	.998
RA	.977	.977	.981	.977	.973

Subject 32	50	25	0	25	50
RS	.927	.948	.988	.951	.949
G	.981	.996	.992	.997	.991
RA	.891	.928	.964	.925	.935

Subject 36	50	25	0	25	50
RS	.981	.997	.997	.995	.995
G	.999	.999	.999	.998	.997
RA	.959	.973	.982	.970	.980

Subject 33	50	25	0	25	50
RS	.981	.994	.996	.996	.985
G	.995	.999	.996	.999	.989
RA	.972	.974	.980	.975	.962

Subject 37	50	25	0	25	50
RS	.972	.986	.980	.994	.984
G	.999	.997	.997	.998	.991
RA	.952	.953	.956	.968	.958

Subject 34	50	25	0	25	50
RS	.917	.982	.996	.993	.988
G	.994	.997	.999	.996	.998
RA	.887	.951	.985	.959	.975

Subject 38	50	25	0	25	50
RS	.992	.994	.997	.997	.996
G	.998	.999	.997	.998	.997
RA	.983	.965	.981	.971	.977

Data Analysis for Group Ax

Subject 1	0	25	50	25	0
RS	.969	.982	.988	.975	.993
G	.990	.995	.994	.999	.999
RA	.951	.933	.970	.953	.975

Subject 2	0	25	50	25	0
RS	.993	.982	.948	.984	.988
G	.997	.998	.997	.998	.998
RA	.965	.943	.942	.943	.978

Subject 3	0	25	50	25	0
RS	.992	.976	.907	.939	.947
G	.995	.995	.934	.956	.997
RA	.962	.931	.830	.869	.944

Subject 4	0	25	50	25	0
RS	.977	.930	.908	.930	.981
G	.999	.992	.987	.979	.999
RA	.962	.855	.879	.868	.969

Subject 5	0	25	50	25	0
RS	.956	.954	.951	.961	.990
G	.995	.996	.989	.992	.998
RA	.947	.922	.918	.959	.966

Subject 6	0	25	50	25	0
RS	.989	.978	.939	.970	.961
G	.999	.981	.989	.990	.999
RA	.970	.936	.911	.921	.956

Subject 7	0	25	50	25	0
RS	.966	.954	.871	.959	.974
G	.999	.955	.980	.994	.997
RA	.932	.890	.830	.948	.940

Subject 8	0	25	50	25	0
RS	.979	.995	.987	.996	.998
G	.994	.991	.985	.993	.997
RA	.961	.954	.952	.956	.971

Subject 9	0	25	50	25	0
RS	.983	.980	.973	.990	.991
G	.999	.992	.973	.999	.996
RA	.966	.947	.933	.961	.976

Subject 10	0	25	50	25	0
RS	.978	.887	.725	.932	.966
G	.993	.985	.962	.983	.995
RA	.960	.809	.687	.886	.941

Data Analysis for Group Bx

Subject 16	50	25	0	25	50	Subject 21	50	25	0	25	50
RS	.944	.863	.987	.775	.512	RS	.821	.966	.777	.964	.774
G	.962	.986	.992	.986	.992	G	.921	.944	.990	.986	.985
RA	.907	.818	.967	.654	.459	RA	.778	.932	.951	.925	.770

Subject 17	50	25	0	25	50	Subject 22	50	25	0	25	50
RS	.812	.981	.983	.963	.719	RS	.670	.903	.952	.860	.725
G	.997	.999	.990	.997	.958	G	.897	.984	.995	.999	.952
RA	.806	.969	.944	.934	.698	RA	.618	.847	.959	.800	.701

Subject 18	50	25	0	25	50	Subject 23	50	25	0	25	50
RS	.734	.929	.939	.883	.817	RS	.987	.996	.997	.995	.994
G	.923	.979	.993	.996	.954	G	.982	.996	.997	.994	.991
RA	.624	.871	.906	.855	.732	RA	.946	.965	.976	.956	.963

Subject 19	50	25	0	25	50	Subject 24	50	25	0	25	50
RS	.932	.986	.983	.966	.967	RS	.966	.961	.973	.973	.874
G	.968	.991	.996	.998	.992	G	.990	.996	.985	.995	.987
RA	.876	.944	.968	.917	.928	RA	.934	.939	.939	.939	.802

Subject 20	50	25	0	25	50	Subject 25	50	25	0	25	50
RS	.989	.995	.995	.995	.992	RS	.574	.759	.944	.547	.165
G	.985	.994	.992	.991	.992	G	.901	.914	.988	.959	.911
RA	.951	.965	.961	.967	.963	RA	.506	.722	.916	.573	.133

APPENDIX III

DATA ANALYSIS

Groups A + B

0		25		50	
1.9459	1.9720	2.0756	2.0286	1.6658	2.2269
2.5147	1.9210	1.7828	1.6437	1.7467	1.9333
2.3507	2.2975	1.9588	2.1648	2.0438	1.8527
2.1847	2.2975	1.7555	1.8421	1.7735	2.3795
1.9459	2.1457	2.1457	2.2269	2.0139	1.9588
2.1273	1.9996	1.9720	2.1457	1.8216	1.6967
1.8634	2.2975	1.8216	1.8634	1.6157	1.9720
2.0595	2.4426	1.8317	2.0139	2.2494	2.1847
2.1095	2.3234	1.8317	2.1648	1.8317	2.1457
1.0687	2.3507	1.8857	1.6225	1.4267	2.2975
2.2729	1.8972	2.0438	2.1847	2.1273	1.9210
1.8116	2.3234	1.8744	1.9333	1.4076	2.2269
		1.7555	2.2269		
		1.7735	2.0922		
		1.7922	2.0595		
		2.2269	2.1095		

Source of Variation	df	SS	MS
Between	2	0.39773997	0.198869985
Within	77	4.61941827	0.0599924451
Total	79	5.01715824	

F ratio = 3.314917153

significant at $F_{0.05}$

Groups A and B

Means	0	25	50
	2.105120833	1.964059375	1.938308333

Means L → H 50 25 0

$$S_{\bar{X}\cdot j} = \sqrt{\frac{\text{Error Mean Square}}{\text{No. of obs in } \bar{X}\cdot j}}$$

$S_{\bar{X}.0} = 0.0499968519$

$S_{\bar{X}.25} = 0.0432985439$

$S_{\bar{X}.50} = 0.0499968519$

At 95% level P = 2 3

2.82 2.97

$2.82(S_{\bar{X}.0}) = 0.1409911224$

$2.82(S_{\bar{X}.25}) = 0.1221018938$

$2.82(S_{\bar{X}.50}) = 0.1409911224$

$2.97(S_{\bar{X}.0}) = 0.148906501$

$2.97(S_{\bar{X}.25}) = 0.1285966754$

$2.97(S_{\bar{X}.50}) = 0.1484906501$

0% - 50% = 0.1668125 > 0.148906501

0% - 25% = 0.141061458 > 0.1409911224

25% - 50% = 0.0257510420 < 0.1409911224

∴ 0% different from 50% and 25%

25%-50% are not different for this data

<u>Groups A and B</u>

$R(RA, G) = 0.659436367$ denote $R(RA, G)$ as $\hat{\rho}_1$

$R(RA, RS) = 0.946386516$ denote $R(RA, RS)$ as $\hat{\rho}_2$

	1	2
$\hat{\rho}_i$	0.659436367	0.946386516
z_i	0.7918156483	1.79596463
$n_i - 3$	77	77
$z_i(n_i - 3)$	60.96980492	138.2892765

$H_0: \rho_1 = \rho_2$

$$z_i = \frac{1}{2} \ln\left(\frac{1+\hat{\rho}_i}{1-\hat{\rho}_i}\right) \qquad \bar{z} = \frac{\sum(n_i - 3) z_i}{\sum(n_i - 3)} = 1.293890139$$

$W = \sum(n_i - 3)(z_i - \bar{z})^2$ distributed as χ^2 1 df.

$(z_1 - \bar{z})^2 = 0.2520787942$

$(z_2 - \bar{z})^2 = 0.2514986161$

$W = 38.7756059$ Reject at 0.0001

Groups Ax and Bx

0		25		50	
1.8421	1.8972	1.6810	2.0756	2.0922	1.0402
2.0139	1.7380	1.7644	1.3372	1.7555	0.7217
1.9720	2.1095	1.6658	2.0139	1.1881	1.7922
1.9206	2.2053	1.2744	1.6734	1.3713	1.6888
1.8018	1.7467	1.6022	1.2454	1.5761	0.5573
2.0922	2.0438	1.7047	2.0139	1.5333	0.4960
1.6734	1.7735	1.4219	1.7295	1.1881	0.8633
1.9588	1.5047	1.8744	0.9118	1.8527	0.9330
2.0286	2.0595	1.8018	0.7822	1.6810	1.6437
1.9459	1.9588	1.1241	1.6888	0.8422	1.9856
2.1847	1.8421	1.8634	1.2744	1.5103	1.0203
2.2494	1.9333	1.7644	1.5698	1.1155	0.8692
1.7735	2.2053	1.3289	2.0438	0.7315	1.9856
2.0756	1.7295	1.9333	1.6225	1.3583	1.1041
2.0286	1.5635	1.5955	1.0986	1.18421	0.1337
		1.8116	1.9210		
		1.8972	1.7735		
		1.9588	1.3249		
		1.4030	1.7295		
		1.1507	0.6519		

Source of Variation	df	SS	MS
Between	2	6.2909971	3.14549855
Within	97	13.28113115	0.13691887
Total	99	19.57212825	

F ratio = 22.97344676 significant at $F_{.01}$

Groups Ax and Bx

Means	0	25	50
	1.92906	1.5775775	1.28243

Means L → H 50 25 0

$$S_{\bar{X}.j} = \sqrt{\frac{\text{Error Mean Square}}{\text{No. of Obs. in } \bar{X}.j}}$$

$S_{\bar{X}.0} = 0.0675571062$
$S_{\bar{X}.25} = 0.0585061698$
$S_{\bar{X}.50} = 0.0675571062$

At 95% Level P = 2 3
2.80 2.95

$2.80(S_{\bar{X}.0}) = 0.1891598974$ $2.95(S_{\bar{X}.0}) = 0.1992934633$
$2.80(S_{\bar{X}.25}) = 0.1638172754$ $2.95(S_{\bar{X}.25}) = 0.1725932009$
$2.80(S_{\bar{X}.50}) = 0.1891598974$ $2.95(S_{\bar{X}.50}) = 0.1992934633$

0% − 50% = 0.64663 > 0.1992934633
0% − 25% = 0.3514825 > 0.1891598974
25% − 50% = 0.29514750 > 0.1891598974

∴ 0% different from 25% and 50%
25% different from 50%

Groups Ax and Bx

$R(RA,G) = 0.705461256$ denote $R(RA,G)$ as $\hat{\rho}_1$

$R(RA,RS) = 0.985904209$ denote $R(RA,RS)$ as $\hat{\rho}_2$

	1	2
$\hat{\rho}_i$	0.705461256	0.985904209
z_i	0.8780901658	2.473976684
$n_i - 3$	97	97
$z_i(n_i - 3)$	85.17474608	239.9757383

$H_0 ; \rho_1 = \rho_2$

$z_i = \frac{1}{2} \ln\left(\frac{1+\hat{\rho}_i}{1-\hat{\rho}_i}\right)$ $\bar{z} = \frac{\Sigma(n_i-3) z_i}{\Sigma(n_i-3)} = 1.676033425$

$W = \Sigma(n_i-3)(z_i-\bar{z})^2$ distributed as χ^2 1df

$(z_1-\bar{z})^2 = 0.636713449$
$(z_2-\bar{z})^2 = 0.6366826428$

$W = 123.5194205$ Reject at 0.0001

APPENDIX IV

MATHEMATICS OF LENS MODEL

Achievement Correlation RA

The contents of Appendix IV is taken from an article by Tucker, (1964), which is based on an article by Hammond, Hursch and Summers, (1964).

To simplify the algebra involved the random variables YE, YS, $X_1 \ldots X_n$ are standardized. The variables $\hat{YE}$ and $\hat{YS}$ are calculated by multiple regression techniques. The multiple regression equations take the following form:

$$\hat{YE} = B_{e,1}X_1 + \ldots + B_{e,N}X_N$$

$$\hat{YS} = B_{x,1}X_1 + \ldots + B_{x,N}X_N$$

From elementary statistics $YE = \hat{YE} + ZE$ and $YS = \hat{YS} + ZS$ where ZE and ZS are residuals and $Var(\hat{YE}) = (RE)^2$; $Var(\hat{YS}) = (RS)^2$. Also it is known from elementary statistics that $Var(ZE) = 1 - (RE)^2$ and $Var(ZS) = 1 - (RS)^2$.

Previously in the text RA was defined as the correlation between YE and YS so

$$RA = Cov(YE\ YS) / \sqrt{Var(YE)}\ \sqrt{Var(YS)}$$

but

$$\mathrm{Var}(YE) = \mathrm{Var}(YS) = 1$$

so

$$RA = \mathrm{Cov}\ (YE\ YS) = E(YE\ YS)$$

but

$$E(YE\ YS) = E\left[(B_{e,1}X_1 + \ .\ .\ .\ + B_{e,N}X_N + ZE)(B_{s,1}X_1 + \ .\ .\ .\ + B_{s,N}X_n + ZS)\right]$$

$$= E\left[\left(\sum_{i,j=1}^{N} B_{e,i}B_{s,j}X_iX_j\right) + E(ZEZS)\right.$$

$$E(YE\ YS) = \sum_{i,j=1}^{N} B_{e,i}B_{s,j}\ E(X_iX_j) + E(ZE\ ZS)$$

From elementary statistics the correlation coefficient ρ is defined as follows:

$$\rho = E(X_1X_2)/\sigma_1\sigma_2$$

To use the results above some new terms will be defined.

Let C be the correlation between the residual values ZE and ZS.

Let G be the correlation between $Y\hat{S}$ and $Y\hat{E}$. Then

$$C\sqrt{\mathrm{Var(ZE)}}\sqrt{\mathrm{Var(ZS)}} = E(ZE\ ZS)$$

Substituting for Var(ZE) and VAR(ZS) the E(ZE ZS) equals the following expression:

$$E(ZE\ ZS) = C\sqrt{1 - (RE)^2}\sqrt{1 - (RS)^2}$$

Now

$$\sum_{i,j=1}^{N} B_{e,i}X_iB_{s,j}X_j = \sum_{i=1}^{N} B_{e,i}X_i \sum_{j=1}^{N} B_{s,j}X_j = Y\hat{E}\ Y\hat{S}$$

so

$$E\Big(\sum_{i,j=1}^{N} B_{e,i}B_{s,j}X_iX_j\Big) = E(Y\hat{E}\ Y\hat{S})$$

$$= G\sqrt{\mathrm{Var}(Y\hat{E})}\sqrt{\mathrm{Var}(Y\hat{S})}$$

$$\therefore \quad \sum_{i,j=1}^{N} B_{e,i} B_{s,j} \; E(X_i X_j) = G \; RE \; RS$$

Hence

$$RA = G \; RE \; RS + C \sqrt{1 - (RE)^2} \sqrt{1 - (RS)^2}$$

Thus

$$C = \frac{RA - G \; RE \; RS}{\sqrt{1 - (RE)^2} \sqrt{1 - (RS)^2}}$$

Now to see if man out performs the model or vice versa we need to define two more terms.

Let RM be the correlation between $Y\hat{S}$ and YE. This means RM is the validity coefficient of the linear model of the subject.

Let Δ be the differential validity of model over man that is $\Delta = RM - RA$.

Substitute RM for RA in

$$RA = G \; RE \; RS + C \sqrt{1 - (RE)^2} \sqrt{1 - (RS)^2}$$

But since RM is $Cov(Y\hat{S}, YE)$ we get $RS = R(Y\hat{S}, Y\hat{S}) = 1$, thus

$$RM = G\ RE$$

$$\therefore \qquad \Delta = G\ RE - (G\ RE\ RS + C\sqrt{1 - (RE)^2}\sqrt{1 - (RS)^2}\ )$$

or

$$\Delta = G\ RE(1 - RS) - C\sqrt{1 - (RE)^2}\sqrt{1 - (RS)^2}$$

Therefore if $\Delta > 0$ then

$$G\ RE(1 - RS) > C\sqrt{1 - (RE)^2}\sqrt{1 - (RS)^2}$$

which means the model out performed man.

If YE is linear then RE = 1 and the right side goes to zero telling us to use the model.

As RS goes to 1 man becomes more predictable by a linear model and thus there is less difference between him and the linear model.

BIBLIOGRAPHY

Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, Inc., 1958.

Bowman, E. H., "Consistency and Optimality in Managerial Decision Making", Management Science, 1963, 9, 310-321.

Brehmer, B., "Inference Behavior in a Situation Where the Cues are Not Reliably Perceived," Organizational Behavior and Human Performance, 1970, 5, 330-347.

Brehmer, B. and Lindberg, C. A., "The Relation Between Cue Dependency and Cue Validity in Single-Cue Probability Learning with Scaled Cue and Criterion Variables," Organizational Behavior and Human Performance, 1970, 5, 542-554.

Brunswik, E., "The Conceptual Framework of Psychology," Chicago University of Chicago Press, 1953 (Int. Encycl. Unified Sci., Vol. 1 No. 10).

Goldberg, L. R., "Simple Models or Simple Processes? Some Research on Clinical Judgments," American Psychologist, 1968, 23, 483-496.

Goldberg, L. R., "Man Versus Model of Man: A Rationale, Plus Some Evidence, For a Method of Improving on Clinical Inferences", Psychological Bulletin, 1970, 73, 422-432.

Goldberg, L. R., "Five Models of Clinical Judgment: An Emperical Comparison Between Linear and Nonlinear Representations of the Human Inferences Process," Organizational Behavior and Human Performance, 1971, 6, 458-479.

Gray, C. W., "Predicting with Intuitive Correlations," Psychonomic Science, 1968, 1, 41-42.

Gray, C. W., Barnes, C. B. and Wilkinson, E. F., "The Process of Prediction as a Function of the Correlation Between Two Scaled Variables" Psychonomic Science, 1965, 3, 231-232.

Graybill, F. A., An Introduction to Linear Statistical Models, Vol. 1, McGraw Hill, 1961.

Gregory, R. L., The Intelligent Eye, New York, McGraw Hill, 1970.

Hammond, K., "Probabilistic Functioning and the Clinical Method," Psychol. Rev., 1955, 62, 255-262.

Hammond, K. R., "Probabilistic Functionalism": Ego Brunswik's integration of the history, theory and method of psychology. In K. R. Hammond (Ed.), The Psychology of Egon Brunswik, New York: Holt, Rinehart and Winston, 1966.

Hammond, K., Hursch, C. J., and Todd, F. J., "Analyzing the Components of Clinical Inference," Psychol. Rev., 1964, 71, 438-456.

Hammond, K. and Summers, D. A., "Cognitive Control," Psychological Review, 1972, 79, 58-67.

Hicks, C. R., Fundamental Concepts in the Design of Experiments, Holt, Reinhart, and Winston, Inc., 1964.

Kunreuther, H., "Extensions of Bowman's Theory on Managerial Decision-Making," Management Science, 1969, 15, 415-439.

Lee, J. C. and Tucker, R. B., "An Investigation of Clinical Judgment: A Study in Method," Journal of Abnormal and Social Psychology, 1962, 64, 272-280.

Miller, I. and Freund, J. E., Probability and Statistics for Engineers, Englewood Cliffs, N. J., Prentice-Hall, Inc., 1965.

Mood, A. M. and Graybill, F. A., Introduction to the Theory of Statistics, McGraw Hill, 1963.

Newton, J. R., "Judgment and Feedback in a Quasi-Clinical Situation," Journal of Personality and Social Psychology, 1965, 1, 336-342.

Slovic an Lichtenstein, "Bayesian and Regression Approaches in Judgement," Organizational Behavior and Human Performance, 1971, 6, 644-744.

Summers, S. A., "The Learning of Response to Multiple Weighted Cues," Journal of Experimental Psychology, 1962, 64, 29-34.

Todd, F. J. and Hammond, K. R., "Differintial Feedback in Two Multiple-Cue Probability Learning Tasks," Behavioral Science, 1965, 10, 429-435.

Tucker, L. R., "A Suggested Alternative Formulation in the Developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch and Todd," Psychological Review, 1964, Vol. 71, 6, 528-530.

Uhl, C. H., "Learning of Interval Concepts: J. Effects of Differences in Stimulus Weights," Journal of Experimental Psychology, 1963, 66, 264-273.